

IN THE UNITED STATES
PATENT AND TRADEMARK OFFICE

Applicant: Stephen R. Lawrence
Application No.: 10/815,150
Filing Date: March 31, 2004
Title: Systems And Methods For Analyzing Boilerplate
Examiner: Navneet K. Ahluwalia
Group Art Unit: 2166
Atty. Dkt. No.: 24207-10085

CERTIFICATE OF ELECTRONIC (EFS-WEB) TRANSMISSION

I hereby certify that this correspondence is being transmitted via the Office electronic filing system in accordance with 37 C.F.R. § 1.8(a)(i)(C) from the **Pacific Time Zone** of the United States on the local date shown below.

Dated: January 15, 2008 By: /Jie Zhang/
Jie Zhang, Reg. No.: 60,242

MAIL STOP APPEAL BRIEF - PATENTS
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

APPEAL BRIEF

Pursuant to the requirements of 37 C.F.R. § 41.37, please consider this document as Appellant's Brief in the present application currently before the Board of Patent Appeals and Interferences (hereinafter "the Board").

I. Real Party in Interest

The subject application is owned by Google Inc., a Delaware corporation having a place of business at 1600 Amphitheatre Parkway, Building 41, Mountain View, CA 94043.

Assignment from Stephen R. Lawrence to Google Inc. was recorded on September 29, 2004 at Reel 015837, Frame 0244.

II. Related Appeals and Interferences

There are no known related appeals or interferences that may directly affect, be directly affected by, or have a bearing on the Board's decision in the pending appeal.

III. Status of Claims

Claims 1-26 stand finally rejected in the Final Office Action mailed May 10, 2007 (hereinafter referred to as "Office Action"). Specifically, claims 1-26 were rejected under 35 USC § 103(a) as being unpatentable over U.S. Patent Application Publication No. 2004/0036716 to Jordahl ("Jordahl") in view of U.S. Patent Application Publication No. 2004/0030741 to Wolton et al. ("Wolton").

The rejection of claims 1-26 is hereby appealed.

IV. Status of Amendments

All claim amendments submitted to the Examiner during prosecution of the present application have been entered. No amendments were proposed subsequent to issuance of the Office Action. The claims involved in the present appeal are presented in Section VIII of this document.

V. Summary of the Claimed Subject Matter

The claimed invention is generally directed to methods and computer program products that identify and analyze boilerplate elements in articles. (See, e.g., Spec. p. 3, lines 7-12). Boilerplate elements include, for example, copyright notices, “Terms of Service” notices, and “Help” links. (See, e.g., Spec. p. 15, lines 15-18). In one embodiment, an indexer 130 identifies a common element in a plurality of related articles. The indexer 130 then classifies the common element as boilerplate. (See, e.g., Spec. p. 14, lines 3-8, and FIG. 1).

Claim 1: Independent claim 1 is a method that comprises: identifying a common element in a plurality of articles (See, e.g., Spec. p. 14, lines 3-5); analyzing a spatial location of the common element in an article of the plurality of articles (See, e.g., Spec. p. 14, lines 14 and 15); and determining whether the common element is a boilerplate element of the article based at least in part on the spatial location (See, e.g., Spec. p. 14, lines 15-18).

Claim 5: Independent claim 5 is a method that comprises: comparing an element in an article to a predetermined list to generate a comparison result (See, e.g., Spec. p. 52, lines 16-17); analyzing a spatial location of the element in the article (See, e.g., Spec. p. 14, lines 14 and 15); and determining whether the element is a boilerplate element of the article based at least in part on the spatial location and the comparison result (See, e.g., Spec. p. 14, lines 15-18).

Claim 8: Independent claim 8 is a method that comprises: identifying a common element in a plurality of articles (See, e.g., Spec. p. 15, lines 3-4); analyzing a link associated with the common element in an article of the plurality of articles (See, e.g., Spec. p. 15, lines 3-11); and determining whether the common element is a boilerplate element of the article based at least in part on the link associated with the common element (See, e.g., Spec. p. 15, lines 11-14).

Claims 9: Dependent claim 9 is a method that incorporates all limitations of independent claim 8 as described above and further comprises: wherein analyzing the link associated with the common element comprises analyzing an address to which the link refers (See, e.g., Spec. p. 15, lines 11-14).

Claims 11: Dependent claim 11 is a method that incorporates all limitations of independent claim 1 as described above and further comprises: responding to the common element being the boilerplate element, removing the boilerplate element from the article (See, e.g., Spec. p. 13, lines 14-17); and indexing the article (See, e.g., Spec. p. 13, lines 17-19).

Claim 14: Independent claim 14 is a tangible computer-readable medium on which is encoded program code (See, e.g., Spec. p. 5, lines 10 and 11, and p. 13, lines 6-9) that comprises: program code for identifying a common element in a plurality of articles (See, e.g., Spec. p. 14, lines 3-5); program code for analyzing a spatial location of the common element in an article of the plurality of articles (See, e.g., Spec. p. 14, lines 14 and 15); and program code for determining whether the common element is a boilerplate element of the article based at least in part on the spatial location (See, e.g., Spec. p. 14, lines 15-18).

Claim 16: Independent claim 16 is a tangible computer-readable medium on which is encoded program code (See, e.g., Spec. p. 5, lines 10 and 11, and p. 13, lines 6-9) that comprises: program code for comparing an element in an article to a predetermined list to generate a comparison result (See, e.g., Spec. p. 52, lines 16-17); program code for analyzing a spatial location of the element in the article (See, e.g., Spec. p. 14, lines 14 and 15); and program code for determining whether the element is a boilerplate element of the article based at least in part on the spatial location and the comparison result (See, e.g., Spec. p. 14, lines 15-18).

Claim 19: Independent claim 19 is a tangible computer-readable medium on which is encoded program code (See, e.g., Spec. p. 5, lines 10 and 11, and p. 13, lines 6-9) that comprises: program code for identifying a common element in a plurality of articles (See, e.g., Spec. p. 15, lines 3-4); program code for analyzing a link associated with the common element in an article of the plurality of articles (See, e.g., Spec. p. 15, lines 3-11); and program code for determining whether the common element is a boilerplate element of the article based at least in part on the link associated with the common element (See, e.g., Spec. p. 15, lines 11-14).

Claim 20: Dependent claim 20 is a tangible computer-readable medium on which is encoded program code (See, e.g., Spec. p. 5, lines 10 and 11, and p. 13, lines 6-9) that incorporates all limitations of independent claim 19 as described above and further comprises: wherein analyzing the link associated with the common element comprises analyzing an address to which the link refers (See, e.g., Spec. p. 15, lines 11-14).

Claim 22: Dependent claim 22 is a tangible computer-readable medium on which is encoded program code (See, e.g., Spec. p. 5, lines 10 and 11, and p. 13, lines 6-9) that incorporates all limitations of independent claim 14 as described above and further comprises: program code for responding to the common element being the boilerplate element, removing the boilerplate element from the article (See, e.g., Spec. p. 13, lines 14-17); and program code for indexing the article (See, e.g., Spec. p. 13, lines 17-19).

VI. Grounds of Rejection to be Reviewed on Appeal

Claims 1-26 were rejected under 35 USC § 103(a) as being unpatentable over Jordahl in view of Wolton. Jordahl and Wolton, either alone or in combination, do not teach each of the limitations in the claimed subject matter. Therefore, this ground of rejection is improper and presented for review.

VII. Argument

A. Claims 1-7, 10-18, and 21-26 are patentable over Jordahl in view of Wolton

To render a claim unpatentable under §103, the prior art reference (or references when combined) must suggest or teach *all* the limitations of the claimed invention. *See In re Royka*, 490 F.2d 981 (C.C.P.A. 1974); 35 U.S.C. § 103(a); MPEP §§ 706.02(j), 2143.03. If even a single claim limitation is not taught or suggested by the prior art, then that claim cannot be rejected under §103 over the prior art. *See In re Glass*, 472 F.2d 1388, 1392 (C.C.P.A. 1973). The Examiner's rejection of claims 1-7, 10-18, and 21-26 is improper because the suggested combination of Jordahl and Wolton does not teach or suggest all of the limitations of the rejected claims.

Specifically, independent claim 1 recites:

A method comprising:
identifying a common element in a plurality of articles;
analyzing a spatial location of the common element in an article of the plurality of articles; and
determining whether the common element is a boilerplate element of the article based at least in part on the spatial location.

Identifying boilerplate is advantageous because it allows boilerplate elements to be distinguished from non-boilerplate (i.e., “content”) elements when performing operations involving the articles.

Jordahl fails to disclose “analyzing a spatial location of the common element in an article of the plurality of articles” and “determining whether the common element is a boilerplate element of the article based at least in part on the spatial location.” Jordahl discloses a system and method for creating an environment where the analysis of similarities and differences

between pieces of information can be customized and displayed in a manner that is easily understood. See Jordahl, Summary, Paragraph [0011]. However, Jordahl neither analyzes spatial locations of elements nor determines whether the elements are boilerplate elements.

The Examiner cited Figure 8 and Paragraphs [0047] and [0048] of Jordahl for the teaching of “analyzing a spatial location of the common element in an article of the plurality of articles; and determining whether the common element is a boilerplate element of the article based at least in part on the spatial location.” Figure 8 and its corresponding description (paragraphs [0141-0143]) describe a system connecting a plurality of databases. Paragraph [0047] of Jordahl teaches about a significance threshold for a category that serves as a cut off line to determine whether an element is a member of the category. Paragraph [0048] of Jordahl teaches about a point of view that includes groups of hierarchically-linked categories. While the cited sections arguably disclose classifying items in categories, they do not teach or suggest analyzing a spatial location of a common element in an article, and determining whether the common element constitutes boilerplate, as is claimed in claim 1. Thus, Jordahl does not teach or suggest the above cited limitations of claim 1.

The Examiner acknowledges that Jordahl does not disclose using the spatial location as claimed. See Office Action, p. 3, paragraph 5. The Examiner asserts that this deficiency is remedied by Wolton. However, this latter reference fails to teach or suggest “determining whether the common element is a boilerplate element of the article based at least in part on the spatial location” as claimed in claim 1.

In contrast to the claimed invention, Wolton discloses a system displaying search results in alternate three-dimensional and two-dimensional graphical visualization formats. See Wolton, Abstract. In Wolton, elements are retrieved from web pages and represented based on their

associative relationship to real world network information sources, the locations of which bear no relationship with their spatial location in articles. See Wolton, paragraphs [0061] and [0527]. The Examiner cited paragraph [0571] as disclosing using the spatial location of an element as a factor on which a boilerplate determination is based. However, the cited paragraph merely suggests that a common element may be represented in different spatial locations in different information environments representing a collection of network locations. See Wolton, p. 23, paragraph [0511]. The Wolton environment representation maps elements within the representation to real world network information sources (e.g., network sources and sites). See Wolton, p. 24, paragraph [0527]. Therefore, the Wolton environment is not an article and an element's spatial location in the environment is not used to determine whether the element is boilerplate in an article.

Accordingly, Appellant respectfully submits that a person of ordinary skill in the art would not find the elements of independent claim 1 obvious in view of the cited references. The rejection of independent claims 5, 14, and 16, and the dependent claims is improper for at least the same reason.

B. Claims 11 and 22 are patentable over Jordahl in view of Wolton

The Examiner's rejection of claims 11 and 22 is improper because the suggested combination of Jordahl and Wolton does not teach or suggest all of the limitations of the rejected claims. In addition to incorporating all limitations of independent claim 1, dependent claim 11 further recites limitations of "responding to the common element being the boilerplate element, removing the boilerplate element from the article; and indexing the article."

The Examiner cited paragraphs [0077], [0096], and [0133] of Jordahl for the teaching of the above-cited limitation. Paragraph [0077] of Jordahl teaches about sequencing items in the

Jordahl system based on data structures of the system and search algorithms. Paragraphs [0096] and [0133] of Jordahl teach that hierarchies of elements in the Jordahl system may be bonded by common elements or structure. These cited sections, like the rest of Jordahl, are totally silent as to removing a boilerplate element from an article and indexing the article. Wolton also does not teach or suggest the limitations of claim 11.

Accordingly, Appellant respectfully submits that a person of ordinary skill in the art would not find the additional elements of dependent claim 11 obvious in view of the cited references. The rejection of dependent claim 22 is improper for at least the same reason.

C. Claims 8, 9, 19, and 20 are patentable over Jordahl in view of Wolton

The Examiner's rejection of claims 8, 9, 19, and 20 is improper because the suggested combination of Jordahl and Wolton does not teach or suggest all of the limitations of the rejected claims. Independent claims 8 and 19 resemble the independent claims discussed above, except that claims 8 and 19 base the boilerplate determination at least in part on an analyzed link associated with the common element.

Specifically, independent claim 8 recites:

A method comprising:

identifying a common element in a plurality of articles;

analyzing a link associated with the common element in an article of the plurality of articles; and

determining whether the common element is a boilerplate element of the article based at least in part on the link associated with the common element.

The claimed invention is advantageous because it allows for efficient identification of boilerplate. For example, links that go to the home page of a site or a page ending in "help.html" or "copyright.html" may be considered boilerplate.

The Examiner rejected independent claim 8 based on paragraphs [0059] and [0099] of Jordahl and paragraph [0571] of Wolton. However, the cited paragraphs of Jordahl merely disclose that categories may be linked together and describe a mechanism to implement the linking. The cited paragraphs of Wolton merely disclose that a common element may be represented in different spatial locations in different information environments. Neither Jordahl nor Wolton disclose or suggest using a link associated with the common element in an article to determine whether the common element is boilerplate.

Accordingly, Appellant respectfully submits that a person of ordinary skill in the art would not find the elements of independent claim 8 obvious in view of the cited references. The rejection of independent claim 19 and the dependent claims is improper for at least the same reason.

D. Claims 9 and 20 are patentable over Jordahl in view of Wolton

The Examiner's rejection of claims 9 and 20 is improper because the suggested combination of Jordahl and Wolton does not teach or suggest all of the limitations of the rejected claims.

In addition to incorporating all limitations of independent claim 8, dependent claims 9 further recites that analyzing the link comprises analyzing an address to which the link refers. The Examiner rejected claim 9 under the same ground given for independent claim 8. However, as argued above with related to claim 8, the cited paragraphs, like the rest of Jordahl and Wolton, fail to disclose or suggest analyzing the address to which the link refers.

Accordingly, Appellant respectfully submits that a person of ordinary skill in the art would not find the elements of dependent claim 9 obvious in view of the cited references. The rejection of dependent claim 20 is improper for at least the same reason.

For the foregoing reasons, Appellant believes that the Examiner's rejections of claims 1-26 were erroneous, and respectfully request reversal.

Respectfully submitted,
STEPHEN R. LAWRENCE

Dated: January 15, 2008

By: /Jie Zhang/

Jie Zhang, Attorney of Record
Registration No. 60,242
FENWICK & WEST LLP
801 California Street
Mountain View, CA 94041
Phone: (650) 335-7297
Fax: (650) 938-5200

VIII. Claims Appendix

The claims involved in the instant appeal are as follows:

1. A method comprising:
identifying a common element in a plurality of articles;
analyzing a spatial location of the common element in an article of the plurality of articles; and
determining whether the common element is a boilerplate element of the article based at least in part on the spatial location.
2. The method of claim 1, further comprising generating an implicit search query including a search term, the search term comprising a term present in a content element of the article, the content element being distinguishable from the boilerplate element.
3. The method of claim 1, wherein the common element comprises a copyright notice.
4. The method of claim 1, wherein the common element comprises a term having a low inverse document frequency measure.
5. A method comprising:
comparing an element in an article to a predetermined list to generate a comparison result;
analyzing a spatial location of the element in the article; and
determining whether the element is a boilerplate element of the article based at least in part on the spatial location and the comparison result.
6. The method of claim 1, wherein analyzing the spatial location of the common element comprises:
determining whether the common element is at the bottom of the article.

7. The method of claim 1, wherein the common element comprises a navigational element of the article.
8. A method comprising:
 - identifying a common element in a plurality of articles;
 - analyzing a link associated with the common element in an article of the plurality of articles; and
 - determining whether the common element is a boilerplate element of the article based at least in part on the link associated with the common element.
9. The method of claim 8, wherein analyzing the link associated with the common element comprises analyzing an address to which the link refers.
10. The method of claim 1, further comprising:
 - analyzing a markup language element proximate to the common element in the article,
 - wherein determining whether the common element is a boilerplate element comprises determining whether the common element is a boilerplate element of the article based at least in part on the markup language element.
11. The method of claim 1, further comprising:
 - responding to the common element being the boilerplate element, removing the boilerplate element from the article; and
 - indexing the article.
12. The method of claim 1, further comprising:
 - determining weights for elements in the article based at least in part on whether the elements are boilerplate elements.
13. The method of claim 12, further comprising:
 - receiving a search query;
 - determining articles relevant to the search query; and

ranking the articles based as least in part on the determined weights.

14. A tangible computer-readable medium on which is encoded program code, the encoded program code comprising:

- program code for identifying a common element in a plurality of articles;
- program code for analyzing a spatial location of the common element in an article of the plurality of articles; and
- program code for determining whether the common element is a boilerplate element of the article based at least in part on the spatial location.

15. The tangible computer-readable medium of claim 14, wherein the encoded program code further comprises:

- program code for generating an implicit search query including a search term, the search term comprising a term present in a content element of the article, the content element being distinguishable from the boilerplate element.

16. A tangible computer-readable medium on which is encoded program code, the encoded program code comprising:

- program code for comparing an element in an article to a predetermined list to generate a comparison result;
- program code for analyzing a spatial location of the element in the article; and
- program code for determining whether the element is a boilerplate element of the article based at least in part on the spatial location and the comparison result.

17. The tangible computer-readable medium of claim 14, wherein analyzing the spatial location of the common element comprises:

- determining whether the common element is at the bottom of the article.

18. The tangible computer-readable medium of claim 14, wherein the common element comprises a navigational element of the article.

19. A tangible computer-readable medium on which is encoded program code, the encoded program code comprising:

- program code for identifying a common element in a plurality of articles;
- program code for analyzing a link associated with the common element in an article of the plurality of articles; and
- program code for determining whether the common element is a boilerplate element of the article based at least in part on the link associated with the common element.

20. The tangible computer-readable medium of claim 19, wherein analyzing the link associated with the common element comprises analyzing an address to which the link refers.

21. The tangible computer-readable medium of claim 14, wherein the encoded program code further comprises:

- program code for analyzing a markup language element proximate to the common element in the article,
- wherein determining whether the common element is a boilerplate element comprises determining whether the common element is a boilerplate element of the article based at least in part on the markup language element.

22. The tangible computer-readable medium of claim 14, wherein the encoded program code further comprises:

- program code for responding to the common element being the boilerplate element, removing the boilerplate element from the article; and
- program code for indexing the article.

23. The tangible computer-readable medium of claim 14, wherein the encoded program code further comprises:
- program code for determining weights for elements in the article based at least in part on whether the elements are boilerplate elements.
24. The tangible computer-readable medium of claim 23, further comprising:
- program code for receiving a search query;
 - program code for determining articles relevant to the search query; and
 - program code for ranking the articles based at least in part on the determined weights.
25. The method of claim 5, wherein the predetermined list comprises terms or phrases, and wherein the comparison result indicates whether the element matches a term or a phrase in the predetermined list.
26. The method of claim 10, wherein the markup language element proximate to the common element comprises a markup language element affecting a display of the common element in the article.

IX. Evidence Appendix

No evidence of the types described in 37 CFR § 41.37(c)(1)(ix) has been submitted during prosecution of the present application.

X. Related Proceedings Appendix

To the best knowledge of Appellant and Appellant's legal representative, there are no decisions rendered by a court or the Board that may directly affect, be affected by, or have a bearing on the decision of the Board in the instant appeal.